The Post-LLM World: Fighting Digital Garbage

GPT-5.2
Gemini 3

January 27, 2026

Abstract

As large language models (LLMs) become cheap and ubiquitous, the cost of producing text, images, and code approaches zero. In parallel, the cost of verifying authenticity, provenance, and usefulness rises sharply. This paper frames the "post-LLM world" as an adversarial information environment dominated by *digital garbage*: low-cost, high-volume artifacts that dilute attention, distort incentives, and degrade knowledge systems. We propose a formal taxonomy of digital garbage, identify "model collapse" feedback loops that threaten data integrity, and outline a toolkit of cryptographic and economic defenses designed to restore signal in a post-trust era.

1 Introduction

Generative models have fundamentally broken the historical link between the quality of an artifact and the effort required to produce it. When anyone can generate plausible, grammatically correct, and aesthetically pleasing content at scale, attention becomes the ultimate scarce resource.

In this environment, "garbage" is not defined by its lack of polish, but by its economic asymmetry. We are entering an era where the internet is no longer a "Global Village," but a "Global Landfill" where the signal is buried under mountains of synthetic debris.

2 Formalizing Digital Garbage

To address the problem, we must define it. We propose that "garbage" is an economic misalignment.

Let $V(a)$ be the expected value of an artifact a to a consumer, $C_p(a)$ be the marginal cost of production for the creator, and $C_v(a)$ be the cost for the consumer to verify the artifact's utility or truth. An artifact a is considered *digital garbage* when

$$C_p(a) \to 0 \quad \text{and} \quad C_v(a) > \mathbb{E}[V(a)]. \tag{1}$$

In other words, digital garbage is any content that is cheaper to produce than it is to evaluate, and whose evaluation cost exceeds its potential benefit. This creates a "tragedy of the information commons," where individual producers gain utility (clicks, SEO, compliance) while the collective information environment is degraded.

3 Taxonomy of Debris

3.1 Recursive Garbage and Model Collapse

One of the most insidious threats is the feedback loop created when LLMs are trained on the outputs of other LLMs. As synthetic content is scraped into future training sets, models begin to lose the

Category	Primary Driver	Core Risk
Synthetic Noise	Arbitrage (SEO/Ads)	Total dilution of search and discovery
Adversarial Garbage	Malice/Influence	Erosion of social and institutional trust
Recursive Garbage	Efficiency/Negligence	"Model collapse" and cultural inbreeding
Instrumental Garbage	Compliance/Process	Bureaucratic bloat and "hallucinated" productivity

"tail" of human distribution—the rare, creative, or nuanced data points.

The result is a convergence toward a confident but low-fidelity "statistical average" that lacks the creative variance of human thought.

3.2 Instrumental Garbage

This represents the internal garbage of organizations. Automated meeting summaries, "refined" emails, and AI-generated status reports create a veneer of productivity. However, if the cost of reading and verifying these documents exceeds their value, the organization accumulates *informational debt* that slows decision-making.

4 Why the Garbage Wins

4.1 Cost Asymmetry

Generating a convincing phishing email or a 2,000-word SEO blog post takes seconds; identifying its flaws or verifying its sources takes minutes or hours of human labor.

4.2 The Collapse of Polish

For decades, "professionalism" (good grammar, high-resolution images) served as a proxy for quality. LLMs have commodified polish, leaving us with no intuitive heuristics for trust.

4.3 Platform Incentives

Current algorithms prioritize engagement. Synthetic content can be hyper-optimized to trigger human outrage or curiosity, out-competing slow, human-verified information.

5 Defenses: Rebalancing the Equation

To fight digital garbage, we must increase the cost of malicious or low-value production and lower the cost of verification.

5.1 Cryptographic Provenance (e.g., C2PA)

We must move from detecting AI (a losing arms race) to proving origin.

High-stakes content (journalism, legal documents) should be cryptographically signed at the point of creation. Identity protocols that verify a human is behind a stream of data will become a "VIP pass" for the post-LLM internet.

5.2 The "Dark Forest" Strategy

As public platforms become saturated with garbage, users retreat to "dark forests": private, unindexed spaces (e.g., Signal, Discord, physical meetups).

The future of discovery may not be a wide-open search engine, but narrow, high-trust pipes where human curators act as the primary filter.

5.3 Verification-Aware Interfaces

UI/UX must evolve to show verification debt. Interfaces can surface (i) chain of custody (where did this come from?) and (ii) source density (three independent human sources versus one synthetic source).

6 Conclusion

The post-LLM world is not a battle of humans against machines, but a battle for the integrity of our shared information environment. If we fail, the internet will not die—it will become a place where signal is so rare that humans stop looking for it.